Solution Structure of a Low Molecular Weight Protein Tyrosine Phosphatase[†,‡]

Timothy M. Logan,[§] Ming-Ming Zhou,[§] David G. Nettesheim,[§] Robert P. Meadows,[§] Robert L. Van Etten,[*,∥] and Stephen W. Fesik[*,§]

Pharmaceutical Discovery Division, Abbott Laboratories, Abbott Park, Illinois 60064, and Department of Chemistry, Purdue University, West Lafayette, Indiana 47907-1393

Received May 3, 1994; Revised Manuscript Received June 21, 1994[⊗]

ABSTRACT: Protein tyrosine phosphatases (PTPs) are important enzymes involved in signal transduction, cell cycle regulation, and the control of differentiation. Despite the importance of this class of enzymes in the control of critical cell processes, very little structural information is available for this family of proteins. In this paper, we present the first solution structure of a protein tyrosine phosphatase. This protein is a low molecular weight cytosolic PTP that was initially isolated from bovine heart. The structure that was determined from 1747 NMR-derived restraints consists of a central four-stranded parallel β-sheet surrounded by four α-helices and a short 3_{10} helix. The phosphate binding site, identified by chemical shift changes upon the addition of the competitive inhibitors phosphate and vanadate, is in a loop region connecting the C-terminal end of the first β-strand with the first α-helix. Residues in the second, fourth, and fifth α-helices and in some of the loop regions connecting the elements of regular secondary structure also contribute to the binding site. The structure determined here is consistent with previous mutagenesis and chemical modification studies conducted on this protein.

Protein tyrosine phosphorylation is an essential component of the intracellular signal transduction pathways that regulate cell growth, proliferation, and differentiation (Fischer et al., 1991; Charbonneau & Tonks, 1992; Walton & Dixon, 1993). The level of protein tyrosine phosphorylation is controlled by the combined activity of two classes of enzymes which catalyze the phosphorylation and dephosphorylation reactions: the protein tyrosine kinases (PTKs)[1] and the protein tyrosine phosphatases (PTPs), respectively. The role of PTKs in activating signaling responses has been well documented (Hunter & Cooper, 1985; Yarden & Ullrich, 1988). However, the structure, substrate affinities, and regulation of the PTPs are poorly characterized.

Like the PTKs, the PTPs include a large number of proteins which appear to be present in all organisms. The PTPs can be classified into different groups according to their physiological roles, subcellular locations, and sequence similarities. One class of PTPs consists of the receptor-like transmembrane PTPs. An example of this class of PTPs is CD45, which is involved in T-cell activation through ligand-modulated dephosphorylation of phosphotyrosine-containing proteins (Tonks et al., 1988; Chiu et al., 1994). The receptor-like PTPs consist of two intracellular catalytic domains and extracellular segments for responding to a variety of ligands. There are also cytosolic PTPs which have a single catalytic domain with sequence similarity to the receptor-like PTPs. These proteins may also contain targeting and regulatory modules such as Src homology 2 (SH2) domains (Shen et al., 1991; Yi et al., 1992; Feng et al., 1993), band 4.1 domains (Yang & Tonks, 1991), and modules homologous to lipid-binding proteins (Gu et al., 1991). The presence of the ligand or protein binding motifs within regulatory domains suggests that the subcellular location of PTPs may play an important role in controlling their activities and substrate specificities. For example, PTPs confined to the nucleus, such as PAC-1 and dPTP61F, may be important in regulating gene expression (Rohan et al., 1993; Zhang & Dixon, 1994).

Within the cytosolic family of protein tyrosine phosphatases is a distinct group of low molecular weight cytoplasmic proteins that are 157 amino acids in length with molecular weights of approximately 18 000. These proteins, found in a variety of tissues including human placenta and bovine heart (Waheed et al., 1988; Zhang & Van Etten, 1990), contain the PTP signature sequence but appeared to exhibit no other sequence homology or similarity to the other PTPs. Generally, these low molecular weight phosphatases exist in two isoforms which are thought to be derived from a single gene by an alternative splicing of the primary RNA transcript (Lazaruk et al., 1993). Although the exact function of these phosphatases is currently unknown, they display *in vitro* phosphatase activity toward a wide variety of phosphotyrosine substrates, including tyrosine-phosphorylated angiotensin, tyrosine kinase p40, erythrocyte band 3, IgG, and EGF receptor (Chernoff & Li, 1985; Boivin & Galand, 1986; Waheed et al., 1988; Ramponi et al., 1989; Zhang & Van Etten, 1990). In addition, it has been shown that overexpression of the bovine low molecular weight PTP in NIH/3T3 fibroblasts inhibits normal and oncogenic v-*erb*B-transformed cell growth, suggesting that this PTP may act on phosphorylated substrates that are activated by growth factor-dependent mitogenic stimulation as well as by oncogenic transformation (Ruggiero et al., 1993).

[†] The research at Purdue University was supported by a grant from the National Institutes of Health (GM 27003).

[‡] The coordinates for the final structures have been deposited with the Brookhaven Protein Data Bank under the filename 1BVH.

* To whom correspondence should be addressed.

[§] Abbott Laboratories.

[∥] Purdue University.

[⊗] Abstract published in *Advance ACS Abstracts,* August 1, 1994.

[1] Abbreviations: PTP, protein tyrosine phosphatase; PTK, protein tyrosine kinase; BHPTP, PTP isolated from bovine heart; PTP1B, PTP isolated from human placenta; pNPP, *p*-nitrophenyl phosphate; 2D, two dimensional; 3D, three dimensional; NOE, nuclear Overhauser effect; NOESY, nuclear Overhauser effect spectroscopy; HMQC, heteronuclear multiple-quantum coherence; HSQC, heteronuclear single-quantum coherence; TOCSY, total correlation spectroscopy; HNHB, three-dimensional experiment which correlates the amide proton and nitrogen with Hβ of the same residue; HNHA-J, three-dimensional experiment that correlates the amide proton and nitrogen with the intraresidue Hα with *J*-modulation of cross-peak intensity; HCCH-TOCSY, three-dimensional experiment which correlates intraresidue protons using magnetization transfer through the attached carbon nuclei; HMQC-J, modified HMQC experiment used to measure the intraresidue HN–Hα couplings.

0006-2960/94/0433-11087$04.50/0

Despite the range of sequence homology between the different protein tyrosine phosphatases, all PTPs share a highly conserved stretch of amino acid residues, known as the signature sequence, consisting of CXXXXXR(S/T)(G/P), where three of the residues preceding the C are hydrophobic residues. Chemical modification (Zhang et al., 1992), ^{31}P NMR spectroscopy (Wo et al., 1992a; Cho et al., 1992), and site-directed mutagenesis (Guan & Dixon, 1991; Cirri et al., 1993; Davis et al., 1994a) studies have demonstrated that the cysteine residue in the signature sequence is the enzymatic nucleophile and that the nearby arginine residue is directly involved in phosphate binding. On the basis of these and other studies, the same catalytic mechanism has been proposed for all PTPs: initial nucleophilic attack on the phosphotyrosine by the active site cysteine followed by nucleophilic displacement of the phosphate by a water molecule (Saini et al., 1981; Zhang & Van Etten, 1991; Cho et al., 1991).

In order to gain a better understanding of the molecular basis for the regulation, chemical mechanism, and substrate specificity of the PTPs, we have undertaken structural studies of a low molecular weight protein tyrosine phosphatase initially isolated from bovine heart (BHPTP). We have recently reported on the backbone ^{1}H, ^{15}N, and ^{13}C chemical shifts and the secondary structure for this protein tyrosine phosphatase (Zhou et al., 1994). In this paper, we present the three-dimensional structure of this protein determined in solution using heteronuclear multidimensional NMR spectroscopy. In addition, the location of the phosphate binding site was identified on the basis of the changes observed in $^{1}H/^{15}N$ HSQC spectra of BHPTP upon binding phosphate or the competitive inhibitor vanadate. The structure is interpreted in terms of previous mutagenesis and chemical modification studies of this enzyme and is compared to the recently reported crystal structure of the catalytic domain of human PTP1B (Barford et al., 1994).

MATERIALS AND METHODS

Sample Preparation. Recombinant BHPTP was expressed in *Escherichia coli* and purified using procedures similar to those described previously (Wo et al., 1992a; Zhou et al., 1994). An additional purification step consisting of reverse-phase HPLC (21.4 × 250 mm Dynamax-300A C18 column with a linear gradient from 32% to 62% acetonitrile in water containing 0.1% TFA) was added to the published procedure to remove trace amounts of protease contamination. The fractions containing the enzyme were combined and concentrated. The protein was denatured in 6 M urea in the presence of 20 mM DTT. Renaturation of the protein was achieved through dialysis against 100 mM sodium acetate buffer containing 30 mM sodium phosphate and 1 mM EDTA at pH 5.0. The renatured protein was further purified using a SP-Sephadex C-50 cation-exchange column as reported previously (Wo et al., 1992a) and concentrated to 1 mM in sodium phosphate buffer (100 mM, pH 5.0) containing 9:1 H_2O:D_2O or 100% D_2O.

NMR Spectroscopy. NMR experiments were performed at 37 °C on Bruker AMX500 and AMX600 NMR spectrometers equipped with triple resonance probes and a pulsed field gradient accessory (AMX600). Pulsed field gradients were used to provide water and artifact suppression using methods described by Bax and Pochapsky (1992). Quadrature detection in the indirectly detected dimensions of all experiments was accomplished using the States–TPPI method (Marion et al., 1989a). All spectra were processed and analyzed on Silicon Graphics workstations using in-house-written software. The complex data in the indirect dimensions were extended by linear prediction prior to Fourier transformation (Olejniczak & Eaton, 1990).

A 3D HCCH-TOCSY (Kay et al., 1993) spectrum was collected on [U-^{15}N,^{13}C]BHPTP in 100% D_2O buffer as 96 × 64 × 1024 complex points using sweep widths of 7246.4 Hz (^{1}H, t_1), 8591.1 Hz (^{13}C, t_2), and 10 000 Hz (^{1}H, t_3). For each complex t_1,t_2 increment, 64 scans were collected for a total measuring time of 5.5 days. The ^{1}H carrier was positioned on the water frequency for both ^{1}H dimensions, and the ^{13}C carrier was placed at 39.8 ppm. Two different experiments with DIPSI-2 (Rucker & Shaka, 1989) mixing times of 9.6 and 19.3 ms were collected under identical spectrometer conditions. In order to assign the aromatic ^{1}H and ^{13}C signals, a separate HCCH-TOCSY experiment was performed with 96 × 33 × 2048 complex points collected over 7575.8 Hz (^{1}H, t_1), 4545.5 Hz (^{13}C, t_2), and 10 000 Hz (^{1}H, t_3) sweep widths, with the ^{13}C carrier positioned at 125.0 ppm. In this experiment, the DIPSI-2 mixing scheme was applied for 6.6 ms, and the total measuring time was 72 h.

A 3D ^{15}N-separated NOESY-HSQC spectrum was collected using an 80-ms mixing time on [U-^{15}N]BHPTP in 90% H_2O. The data were collected as 96 × 64 × 1024 complex points over sweep widths of 7142.9 Hz (^{1}H, t_1), 1947.8 Hz (^{15}N, t_2), and 10 000 Hz (^{1}H, t_3) in 64 scans per complex t_1,t_2 time point for a total experiment time of approximately 120 h. Additional suppression of the residual solvent signal was accomplished by convolution of the time–domain data (Marion et al., 1989b). In addition, a 3D ^{13}C-separated HMQC-NOESY experiment was performed with an 80-ms mixing time on [U-^{15}N,^{13}C]BHPTP in 100% D_2O buffer with 128 × 48 × 1024 complex points over sweep widths of 6578.9 Hz (^{1}H, t_1), 5611.7 Hz (^{13}C, t_2), and 10 000 Hz (^{1}H, t_3), respectively. For each complex t_1,t_2 increment, 64 scans were collected for a total acquisition time of approximately 120 h. The carrier was on the residual HOD signal in both ^{1}H dimensions and was at 66.5 ppm in ^{13}C. Final data set sizes were 1024 (^{1}H, ω_3) × 512 (^{1}H, ω_2) × 128 (^{13}C, ω_1) real points.

A 3D HNHB (Archer et al., 1991) experiment was performed on the AMX500 using [U-^{15}N]BHPTP in 90% H_2O. The data were collected as 64 × 56 × 1024 complex points using sweep widths of 5681.8 Hz (^{1}H, t_1), 1623.4 Hz (^{15}N, t_2), and 8333.33 Hz (^{1}H, t_3) with the ^{1}H carrier at the water frequency and ^{15}N at 116.3 ppm. The total acquisition time was 70 h. Solvent suppression was afforded by spin lock purge pulses (Messerle et al., 1990). Removal of the residual solvent signal was accomplished by convolution of the time–domain data (Marion et al., 1989b).

Stereospecific assignments for the methyl groups in valine and leucine residues were obtained from an analysis of the ^{13}C–^{13}C one-bond couplings observed in a high-resolution ^{1}H, ^{13}C HSQC experiment (Bodenhausen & Ruben, 1980) collected on a nonrandomly ^{13}C-labeled BHPTP sample using the approach described by Wüthrich and co-workers (Neri et al., 1991). The data were collected on the AMX500 spectrometer as 576 (^{13}C, t_1) × 2048 (^{1}H, t_2) complex points yielding acquisition times of 245 and 109 ms for ^{1}H and ^{13}C, respectively.

Structure Calculations. Structure calculations based upon the experimentally derived restraints were performed using X-PLOR following the distance geometry/simulated annealing protocol as described in the X-PLOR 3.1 manual (Brünger, 1992). The starting structures obtained from distance geometry were refined by dynamical simulated annealing following standard protocols (Nilges et al., 1988). The

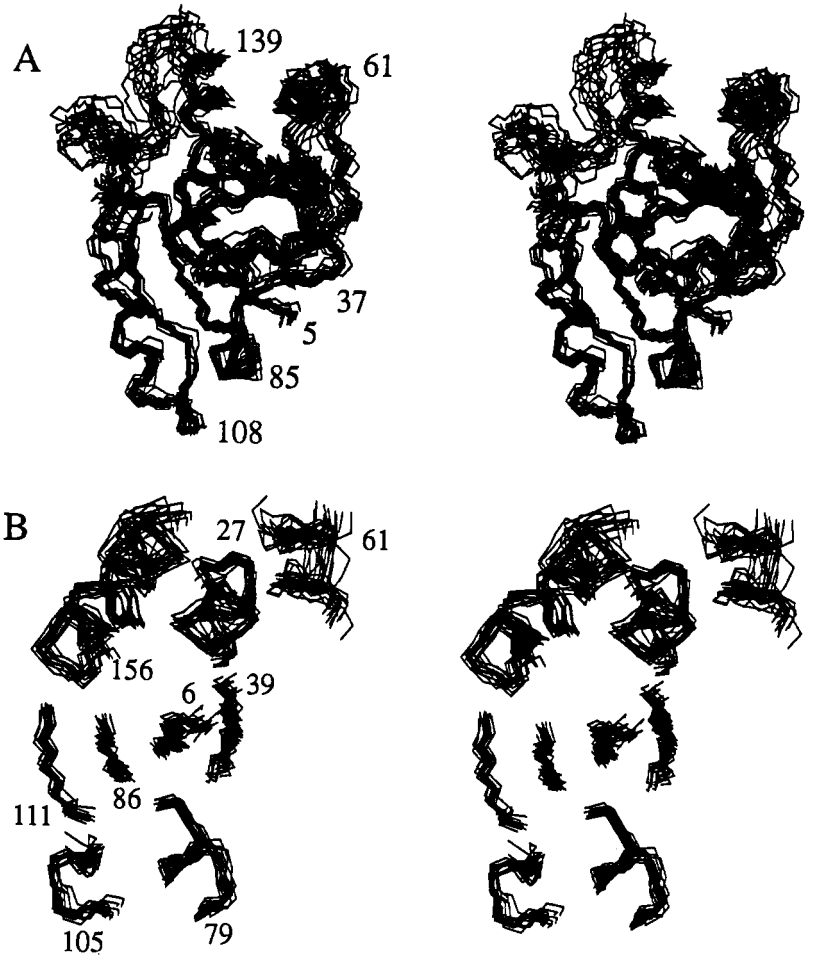

FIGURE 1: Solution structure of BHPTP. Stereoviews showing the superposition of the Cα, C′, and N backbone atoms for residues 5–157 in the 15 lowest energy structures of BHPTP (A) and those residues forming the secondary structure elements (B).

X-PLOR F_{repel} function was used to simulate van der Waals interactions with atomic radii set to 0.75 times their CHARMM values (Brooks et al., 1983). NOE distance restraints were applied with a square-well potential and a force constant of 50 kcal·mol^{-1}·Å^{-2}. Torsional restraints were applied to ϕ angles with a 200 kcal·mol^{-1}·rad^{-2} force constant. Dihedral bounds of $-120 \pm 30°$ were applied for residues with $^3J_{HNH\alpha}$ coupling constants >9.0 Hz and $-60 \pm 30°$ for coupling constants <5.5 Hz. The latter restraints were only applied in regions identified as α-helical from NOE patterns. Hydrogen bonds, identified for slowly exchanging amides based upon initial structures, were given bounds of 1.8–2.3 Å (H→O) and 2.5–3.3 Å (N→O) and applied with a force constant of 50 kcal·mol^{-1}·Å^{-2}. Distance restraints were obtained from the NOE cross-peak intensities and classified as strong, medium, or weak, corresponding to distance bounds of 1.8–3.0, 1.8–4.0, and 1.8–5.0 Å, respectively. Corrections for center averaging (Wüthrich et al., 1983) were applied where appropriate to reflect the proton type (CH_3, CH_2, or aromatic) and the availability of stereospecific assignments.

Inhibitor Titration Experiments. The buffer conditions (100 mM phosphate) used to determine the BHPTP structure result in inhibition of the BHPTP-catalyzed hydrolysis of *p*-nitrophenyl phosphate (pNPP). The phosphate binding site was characterized by comparing the chemical shift changes observed for the backbone amide resonances of BHPTP in ^{1}H, ^{15}N HSQC spectra (Bodenhausen & Ruben, 1980) collected in 100 mM sodium phosphate versus 100 mM sodium acetate (pH 5.0) buffer. Chemical shift changes upon binding sodium vanadate (VO_4^{3-}), a competitive inhibitor for pNPP hydrolysis (K_i = 35 μM; Zhou et al., 1993), were determined from a series of six HSQC spectra collected in 100 mM sodium acetate buffer with VO_4^{3-} concentrations ranging from 0 to 2.5 mM (saturating) levels. The HSQC spectra were acquired as 96 × 1024 complex points using sweep widths of 1947.8 Hz (^{15}N, t_1) and 10 000 Hz (^{1}H, t_2), respectively.

RESULTS AND DISCUSSION

Assignments. Sodium phosphate acts as a competitive inhibitor for the enzymatic activity of BHPTP with a K_i of ~1 mM. Thus, the assignments were obtained and the structure determination was performed on the inhibited, phosphate-bound form of BHPTP. The sequence-specific ^{1}H, ^{13}C, and ^{15}N backbone chemical shift assignments for BHPTP were obtained from a combination of double- and triple-resonance multidimensional NMR experiments as recently described (Zhou et al., 1994). From these chemical shift assignments, we were able to determine the secondary structure of BHPTP. In order to determine the three-dimensional structure of the protein, the side-chain chemical shift assignments were obtained from 3D HCCH-TOCSY experiments by correlating the previously assigned Hα and Cα signals with the remainder of the side chain. Two data sets, collected with TOCSY mixing times of 9.6 and 19.3 ms, provided a complete set of correlations for most aliphatic side chains, allowing most of the ^{1}H and ^{13}C signals to be assigned. The assignment of some of the β-protons could not be made in the HCCH-TOCSY spectra due to missing correlations or spectral overlap. Some of these Hβ assignments were obtained from a 3D HNHB experiment, and all assignments were subse-

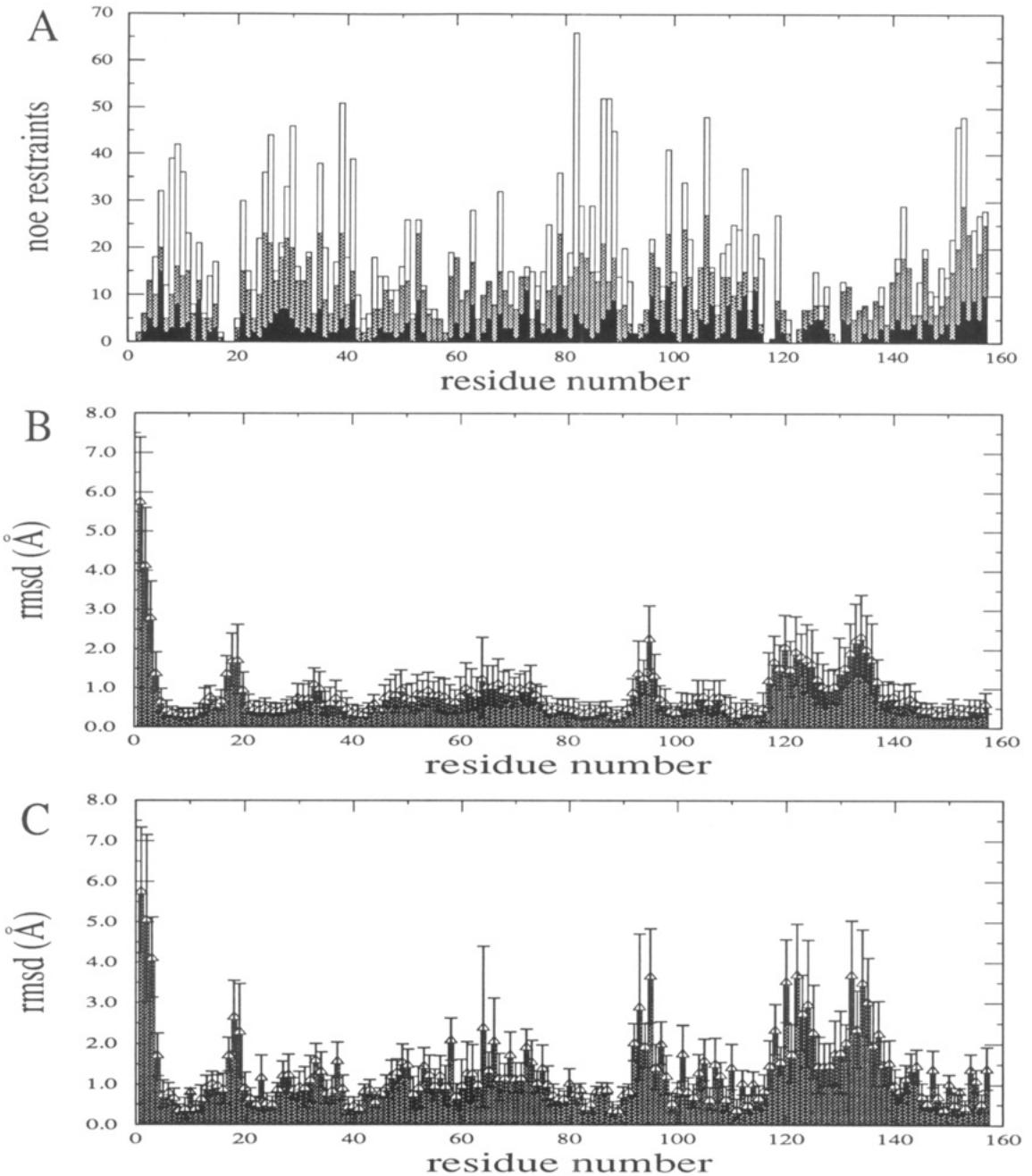

FIGURE 2: (A) Histogram of NOE restraints vs residue number. The height reflects the total number of NOE-derived restraints for each residue while the black, gray, and white portions of the box correspond to the intraresidue restraints, the sum of sequential and short-range restraints, and the long-range restraints, respectively. Residue-based root mean square deviations (Å) of the atomic coordinates from the average structure for (B) the backbone heavy atoms and (C) all non-hydrogen atoms.

quently confirmed by the ^{13}C NOESY-HMQC data. Due to problems of spectral overlap, the side chains of 3 of the 12 arginine spin systems (R58, R64, and R101) could not be assigned. Additional problems attributed to line broadening of the NMR signals precluded the complete assignment of R18, S19, E93, G117, Q122, K123, I127, and P130.

The aromatic spin systems of Phe, Tyr, Trp, and His residues were identified and assigned using data from a 3D HCCH-TOCSY experiment which had been optimized for aromatic $^{1}H/^{13}C$ correlations. Sequence-specific assignment of the aromatic spin systems was accomplished using both ^{15}N- and ^{13}C-separated NOESY data. Despite the combination of these different experiments, Hξ/Cξ assignments for three Phe residues (F10, F26, and F138) could not be made.

Stereospecific assignments of Val and Leu methyl groups were obtained from a 2D ^{1}H, ^{13}C HSQC spectrum collected on a nonrandomly ^{13}C-labeled sample. In this experiment, assignment of the *pro-R* and *pro-S* methyl groups is based on the observation of ^{13}C–^{13}C couplings arising from the stereospecific transfer of the *pro-S* methyl group in the branched chain amino acid biosynthetic pathway (Neri et al., 1989). From these data, stereospecific methyl group assignments were obtained for 7 of 11 leucines and for 11 of 15 valines. Stereospecific assignments could not be determined for all of these residues due to overlapping resonances in either the 2D or 3D experiments.

Structure Determination. Analysis of 3D ^{15}N-separated NOESY-HSQC and ^{13}C-separated HMQC-NOESY experiments provided a total of 1626 approximate interproton distance restraints. Of the restraints used in the structure calculations, 523 were nontrivial intraresidue, 367 were sequential, 221 were between residues separated by less than

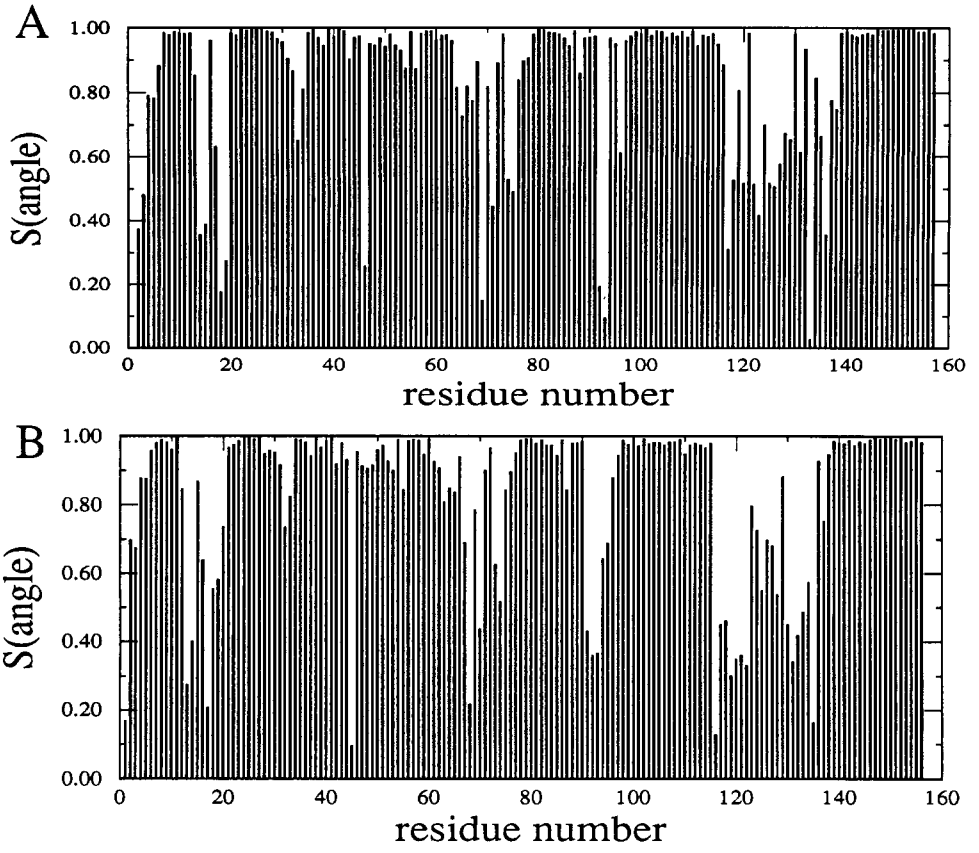

FIGURE 3: Plots of residue-based angular order parameters, S^{angle}, for (A) ϕ and (B) ψ dihedral angles.

Table 1: Structural Statistics and RMSD for 15 PTP Structures[a]

structural statistics	⟨SA⟩	$\langle\overline{SA}\rangle_r$
X-PLOR energies (kcal·mol)		
E_{tot}	170 ± 5	176
E_{repel}[b]	15 ± 2	20
E_{cdih}[c]	0.1 ± 0.1	0.1
E_{NOE}[d]	11 ± 1	7
RMSD from idealized values		
bonds (Å)	0.03 ± 0.00	0.03
angles (deg)	0.40 ± 0.01	0.41
impropers (deg)	0.37 ± 0.02	0.38
Cartesian coordinate RMSD (Å)	N, Cα, C′	heavy atoms
⟨SA⟩ vs $\langle\overline{SA}\rangle$[e]	1.03 ± 0.1	1.60 ± 0.1

[a] ⟨SA⟩ is the ensemble of 15 final X-PLOR structures generated from the DG/SA protocol as described in the X-PLOR 3.1 manual (Brünger, 1992); ⟨SA⟩ is the Cartesian coordinates obtained by averaging ⟨SA⟩ following a least squares superposition of the backbone heavy atoms (N, Cα, C′) for residues 5–157; $\langle\overline{SA}\rangle_r$ is the energy-minimized average Cartesian coordinates. [b] The X-PLOR F_{repel} function was used to simulate van der Waals interactions with atomic radii set to 0.75 times their CHARMM (Brooks et al., 1983) values. [c] Torsional restraints were applied to 61 ϕ angles with bounds of −120 ± 30° for those angles with $^3J_{HN,HA}$ coupling constants >9.0 Hz and −60 ± 30° for coupling constants <5.5 Hz. Restraints for the latter were applied only in the α-helical regions. Force constants of 200 kcal·mol^{-1}·rad^{-2} were employed on all torsional restraints. [d] A total of 1626 nontrivial NOE-derived distance restraints were applied with a square well potential and a force constant of 50 kcal·mol^{-1}·Å^{-2}. In addition, 60 hydrogen bonds were included and given distance bounds of 1.8–2.3 (H→O) and 2.5–3.3 (N→O) Å. No distance restraint was violated by >0.30 Å in any of the final structures. [e] RMSD for residues 5–157.

five residues in the primary amino acid sequence, and 515 were between residues farther apart in the sequence. Dihedral angle (ϕ) restraints for 25 residues that were previously obtained from $^3J_{HNH\alpha}$ coupling constants (Zhou et al., 1994) measured in HMQC-J (Kay & Bax, 1990) and HNHA-J (Vuister & Bax, 1993) experiments were also included in the structure calculations. Additional distance restraints in the form of 30 hydrogen bonds were included in the final structures. The amides participating in hydrogen bonds were identified by slowed amide exchange (Zhou et al., 1994), and the hydrogen bond acceptors were identified from an analysis of initial structures calculated using only NOE and dihedral restraints.

Figure 1A depicts a stereoview of the 15 lowest energy structures of BHPTP that were generated from the NMR-derived restraints using a distance geometry/simulated annealing protocol. As shown in Table 1, the structures satisfy the distance restraints with no violations greater than 0.3 Å and show good covalent geometry and nonbonded contacts. For the majority of the protein, the structure is well-defined by the NMR data with an atomic root mean square (rms) distribution from the mean coordinate positions of 1.03 Å for the backbone atoms and 1.60 Å for all heavy atoms for residues 5–157. When residues L116–E139 (the least defined region of the protein) are excluded, the rms distribution decreases to 0.81 and 1.33 Å for backbone and all heavy atoms, respectively.

As shown in Figure 2, the precision of the NMR-derived structures displayed as the residue-based rms distribution varies greatly for different regions of the molecule and correlates with the number of experimentally determined restraints. The regions of BHPTP with the lowest rms distribution are defined by the largest number of NOEs (Figures 1B and 2), whereas the regions with the higher rms distributions were defined by a smaller number of NOEs. In BHPTP, the residues with the highest rms distributions are found in loop regions. For some residues (V46–S54), short-range and intraresidue NOEs, but no medium- or long-range

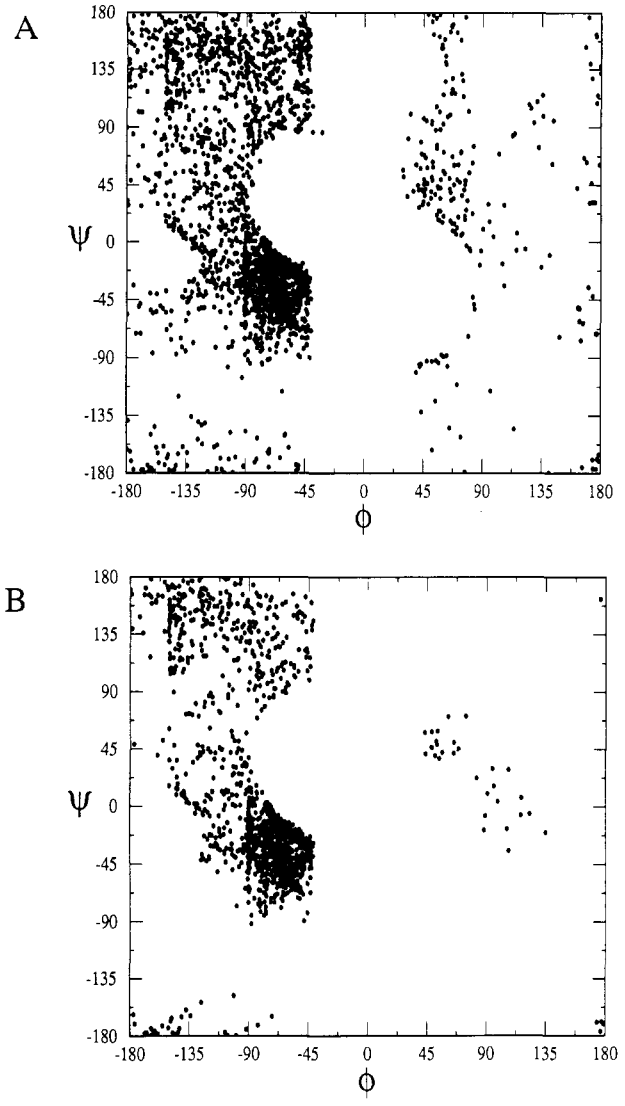

FIGURE 4: Ramachandran plots for (A) all ϕ, ψ angles for the 15 final structures and (B) for all residues in the final structures with ϕ and ψ $S^{angle} > 0.9$.

NOEs, were observed which prevented us from defining the locations of these residues with respect to the remainder of the protein. For other residues, e.g., C17–S19, E93–N95, and D119–I127, the lack of NOEs was attributed to line broadening due to chemical exchange.

Another measure of the precision for the family of structures calculated from the NMR data is provided by the angular order parameter, S^{angle} (Hyberts et al., 1992). S^{angle} is the normalized vector sum of a particular torsion angle over all of the structures. Values approaching zero correspond to an essentially random distribution of angles, whereas S^{angle} values near unity indicate torsional convergence. The S^{angle} values for both ϕ and ψ derived for each residue using the 15 lowest energy BHPTP structures are shown in Figure 3. The angular order parameters indicate that the torsional precision in the structures varies considerably along the backbone of BHPTP, and this variation is in strong agreement with the data presented in Figure 2. As observed with other proteins, the highest value of S^{angle} are associated with residues found in regular secondary structures, which are defined by the largest number of restraints (Figure 2), and the lowest S^{angle} values are observed for residues in regions with little regular secondary structure, such as L116–E139. As shown in Figure 4, Ramachandran plots for the backbone angles of all 15 structures indicate that most of the backbone dihedral angles lie within the energetically favorable regions of ϕ–ψ space (Figure 4), especially for those residues with $S^{angle} > 0.90$ (87 of 155 residues, excluding the N- and C-termini) (Figure 4B).

3D Structure of BHPTP. As shown in Figure 5, the structure of BHPTP consists of a four-stranded twisted parallel β-sheet, four α-helices, one 3_{10} helix, and several connecting loops. The β-sheet is formed by residues K6–V11 (β1), W39–G44 (β2), Y87–C90 (β3), and K112–L115 (β4), with an overall strand folding topology of $+1x, -2x, -1x$, as previously reported (Zhou et al., 1994). Strands β1 and β2 are connected by a loop and an amphipathic α-helix (residues P20–V30, α1). NOEs between residues A22/F10, F26/V8, F26/I41, and V30/I41 indicate that α1 lies on top of the β-sheet and parallel to the strands, forming a right-handed $\beta\alpha\beta$ motif as is commonly observed in many proteins (Rao & Rossmann, 1973; Sternberg & Thornton, 1977). A similar topology is observed for strands β3 and β4, which are connected by the fourth helix (L96–K102, α4) as evidenced by NOEs observed between residues I88/L99, C90/L99, I113/L96, I113/L99, and L115/L96. The right-handedness of this second $\beta\alpha\beta$ motif places α4 on the opposite face of the β-sheet relative to α1.

Between strands β2 and β3 are an α-helix (P57–H66, α2) and two turns of a 3_{10} helix (K79–T84, α3) that are connected by stretches of nonregular secondary structure. The basis for assigning α3 as a 3_{10} helix is the absence of αN (i, i+4) NOEs in the ^{15}N-separated NOESY-HSQC spectrum (Zhou et al., 1994) and the absence of strong $\alpha\beta$ (i, i+3) NOEs in the ^{13}C-separated HMQC-NOESY data. However, the dihedral angles observed in the final structures are consistent with either a 3_{10} or an α-helix. NOEs observed between hydrophobic residues in these two helices orient these helices with respect to the rest of the protein. Most of the amino acid sequence changes associated with the various isoforms of the low molecular weight PTPs (Taga & Van Etten, 1982; Camici et al., 1989; Dissing et al., 1991; Manao et al., 1992; Wo et al., 1992a,b) are located in this region. Changes in the amino acid sequence resulting from nonconserved substitutions can easily be accommodated at this site, and changes in the primary sequence of this loop will result in only minor perturbations to the overall structure of the various low molecular weight PTPs.

At the C-terminus of BHPTP, an α-helix (E139–V156, α5) is found that is connected to β4 by a long loop with little defined secondary structure (L116–F138). Several NOEs were observed between residues in α5 and residues in α1, β1, and β3, indicating that α5 is located next to α1 on the top face of the β-sheet. In all of the final structures, an irregularity in the helicity is observed for α5. In contrast to the ϕ angles of –65° which are typically observed for residues in regular α-helices, a ϕ value of approximately –120° was observed for C149.

Mapping the Inhibitor Binding Site. Sodium phosphate acts as a competitive inhibitor of the BHPTP-catalyzed hydrolysis of *p*-nitrophenyl phosphate (pNPP), a nonphysiological phosphotyrosine analogue. Under the conditions used in the current study, BHPTP is in the inhibited, phosphate-bound form. The inhibitor binding site was investigated by comparing the amide ^{1}H and ^{15}N chemical shifts observed in HSQC spectra of [U-^{15}N]BHPTP collected in the presence (100 mM phosphate buffer, pH 5.0) and absence (100 mM sodium acetate buffer, pH 5.0) of phosphate. Two types of effects were observed in acetate versus phosphate buffer:

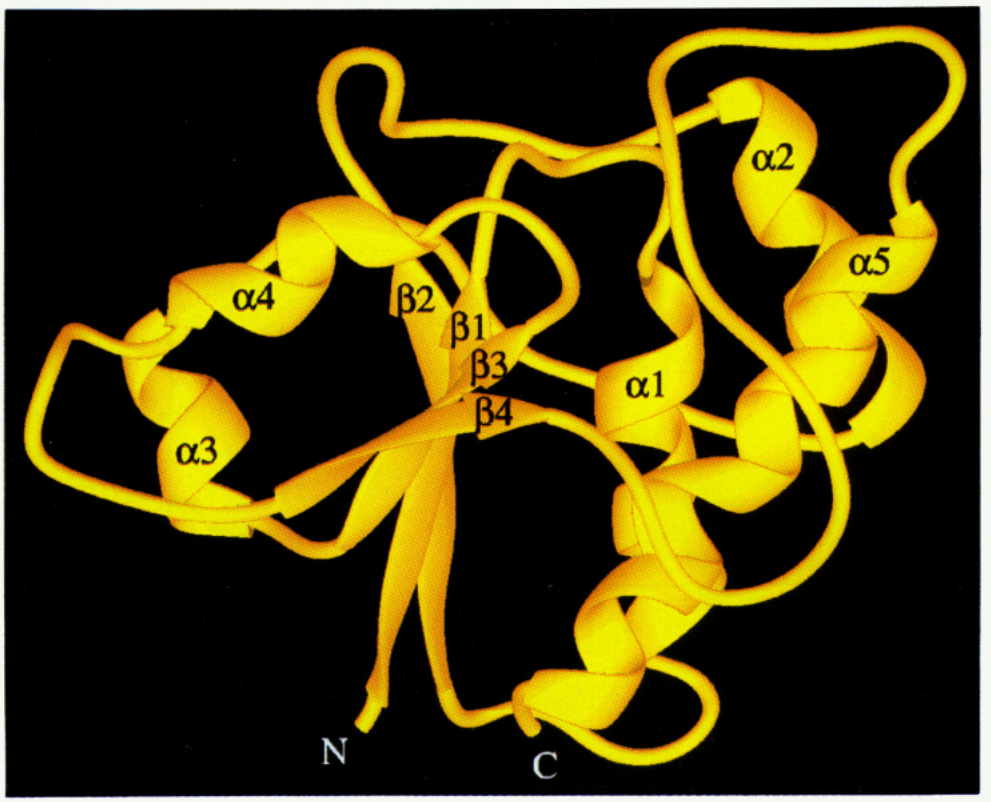

FIGURE 5: Ribbon diagram (Carson, 1987) of the BHPTP structure. The β-sheet strands and helices are labeled according to the discussion in the text.

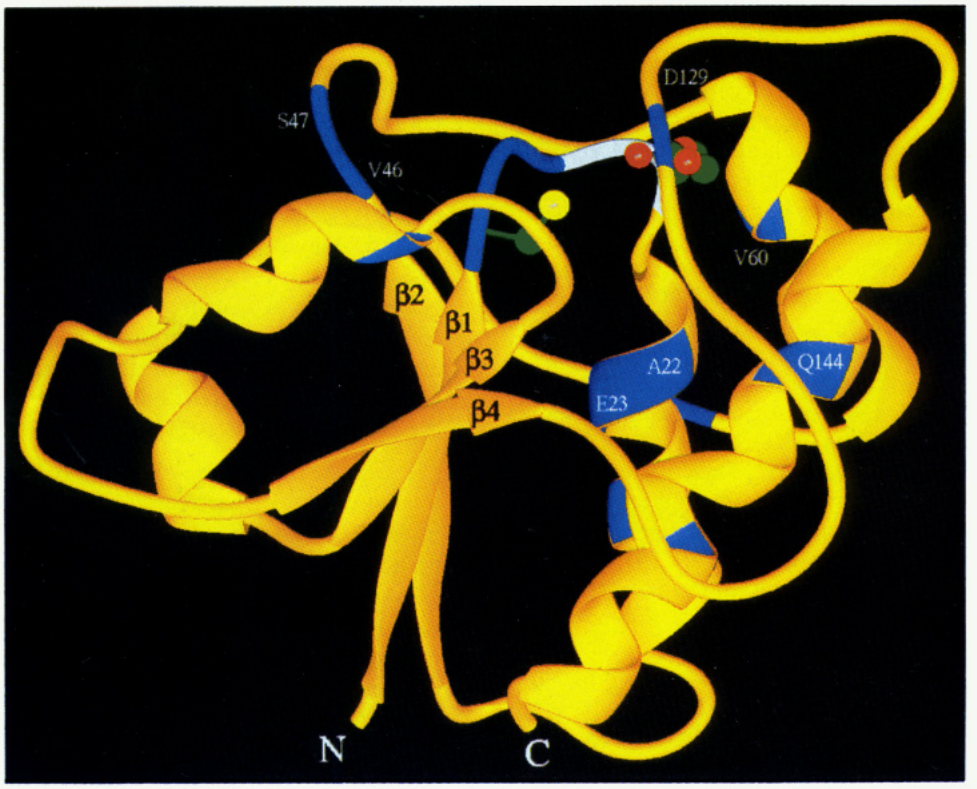

FIGURE 6: Ribbon diagram of BHPTP highlighting the residues forming the active site. The backbone heavy atoms of the residues that exhibit the largest ^{15}N (>|0.27| ppm) and ^{1}H (>|0.04| ppm) chemical shift changes upon binding to the inhibitors are indicated in blue. Residues L16, G17, and R18 are colored in white. The heavy atoms of residues C12 and A18 are shown. The sulfur atom of C12 is colored in yellow; the guanidinium nitrogens of A18 are in red.

altered chemical shifts (residues C12, A22, E23, F26, V46, S47, V60, N69, R75, L96, D129, Q144, and R150) or missing peaks in the $^{15}N/^{1}H$ correlation spectrum (residues L13–S19).

These results were extended by analyzing the effects of added sodium vanadate (VO_4^{3-}), a potential transition-state analogue for phosphatases (Van Etten et al., 1974). Like phosphate, vanadate is a competitive inhibitor of pNPP

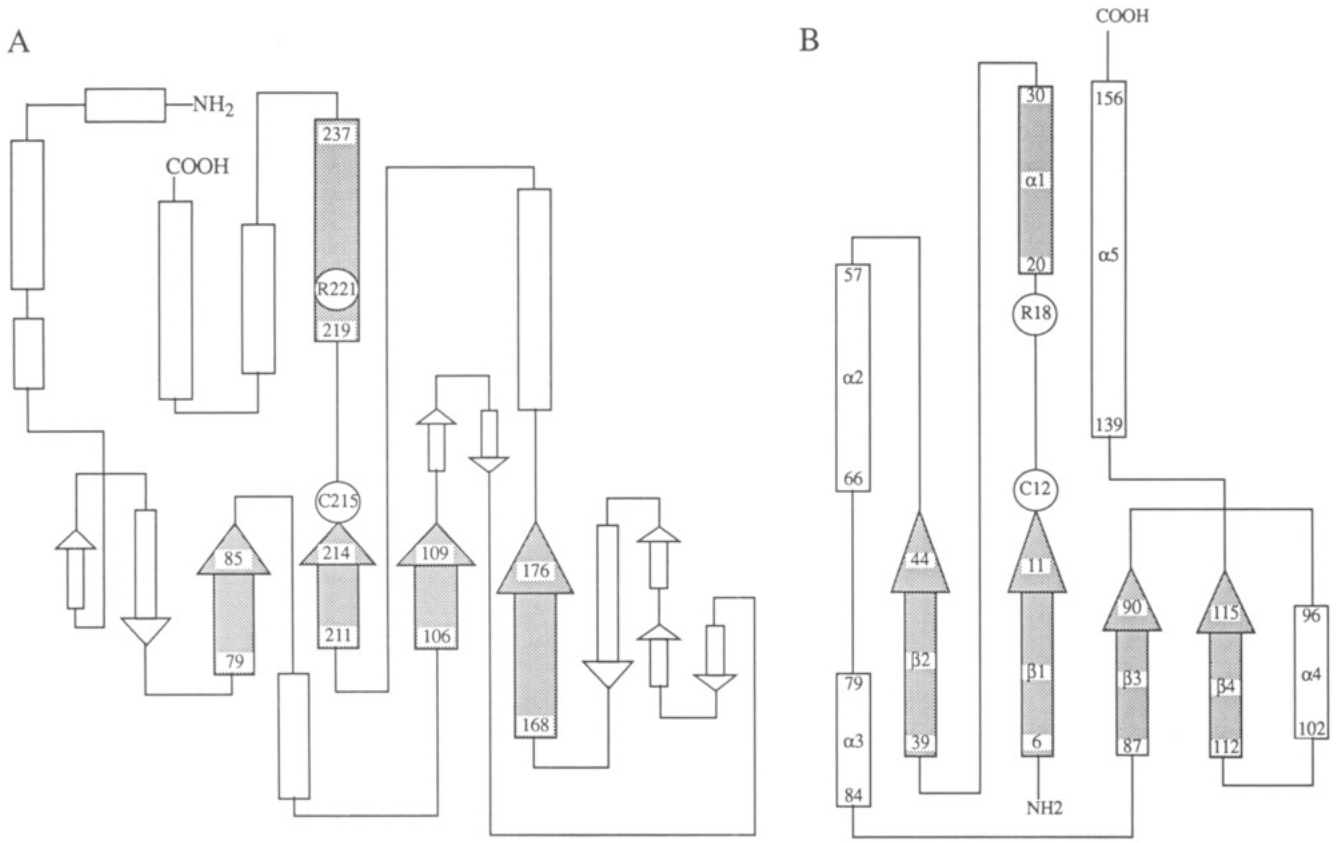

FIGURE 7: Schematic folding topology diagram comparing the secondary structural elements between (A) PTP1B and (B) BHPTP. Arrows represent β-strands, rectangles represent α-helices, and lines indicate regions of nonregular secondary structure. Figure 7A is adapted from Barford et al. (1994).

hydrolysis but exhibits a higher affinity for the active site (K_i = 35 μM; Zhou et al., 1993). It was thought that the tighter binding of VO_4^{3-} compared to phosphate would alter the position of the equilibrium resulting from the chemical exchange such that the residues which exhibited line broadening in phosphate and acetate buffer would be observed. In the absence of VO_4^{3-}, L13–S19 were broadened beyond observation. At intermediate VO_4^{3-} concentrations, correlations for L13–I16 appeared which subsequently sharpened and approached the chemical shifts observed in 100 mM phosphate as the titration continued. At saturating VO_4^{3-} concentrations, all of the residues which were perturbed in phosphate binding were also perturbed upon binding VO_4^{3-} except for C17, R18, and S19, which were not observed. Thus, a similar set of residues is perturbed upon binding both inhibitors. Most of these residues are located at the C-terminal end of one of the central β-strands, with residues in α1 (A22, E23), α2 (V60), α4 (L96), α5 (Q144, R150), and some of the interconnecting loop regions also contributing to the binding (Figure 6). The differences in line broadening are centered around R18, suggesting a critical role for this residue in binding these inhibitors.

The results obtained in this study can be used to explain and extend earlier results on the chemical mechanism of BHPTP. Chemical modification (Zhang et al., 1992), spectroscopic (Wo et al., 1992a), and mutagenesis (Cirri et al., 1993; Davis et al., 1994a) studies have indicated that C12 is the active site nucleophile and that R18 is required for catalytic activity (Cirri et al., 1993; Davis et al., 1994a). Our titration data also demonstrate the involvement of residues C12–R18 in the binding of phosphate. In the three-dimensional structures presented here, R18 is in a loop extending from the C-terminus of β1 into the N-terminus of α1, an arrangement which would allow interactions between both residues and the phosphate group of the phosphotyrosine. However, the NMR data is ambiguous regarding the beginning of α1. Slow amide exchange for residues I21 and A22 (Zhou et al., 1994) indicate extension of α1 at least to S19 and possibly to R18. However, the low signal to noise observed for the residues in this region precluded the identification of NOEs characteristic of α-helices. The location of the phosphate group near the amino terminus of an α-helix would allow additional stabilization of the phosphate-bound form by electrostatic interactions between the phosphate group and the helix macrodipole (Hol et al., 1978).

Previous studies have shown that H72 is involved in the binding and affects the hydrolysis of phosphorylated substrates by BHPTP (Zhou, 1993). H72 has a high pK_a (9.2) which may be due to interactions with one or more nearby anions (Davis et al., 1994b). Kinetic analysis of an H72A mutant revealed that this residue is not directly involved in the chemical mechanism of BHPTP but was postulated to be involved in defining the integrity of the enzyme active site. The present study shows that H72 is located in the loop connecting α2 and α3, positioned away from C12 and the active site. Although the current NMR structures are not able to uniquely orient the H72 side chain, the side chains of E23 and D42 lie above the imidazole group of H72. The amide of E23, but not D42, is perturbed upon binding VO_4^{3-}. Thus, H72 could exert its effects on substrate binding at the active site through electrostatic interactions with the side chain of E23, which is located near the N-terminus of α1. Mutations which destroy this interaction (i.e., H72A) may alter the size and possibly the orientation of this side of the active site.

Comparison with Other Protein Structures. Identification of the folding topology and the location of the active site of the BHPTP structure indicates that this protein belongs to a family of proteins with open α/β structures (Bränden, 1980). Other proteins exhibiting this open α/β folding motif include lactate dehydrogenase (Adams et al., 1970), phosphoglycerate mutase (Campbell et al., 1974), both domains of the arabinose-binding protein (Gilliland et al., 1981), adenylate kinase (Dreusicke et al., 1988), the bacterial chemotaxis response regulator protein Che Y (Stock et al., 1993), flavodoxin (Smith et al., 1977; Stockman et al., 1990), enzyme $IIB^{cellobiose}$ of the sugar transport system (Ab et al., 1994), and rhodanese, a sulfur transferase (Ploegman et al., 1978). The sizes, amino acid sequences, chemistry, and folding topologies of the parallel β-sheet of these proteins are very different from one another. Nevertheless, each of these proteins are involved in binding to phosphorylated (or sulfated) substrates or cofactors, and the ligand binding site is located at the C-terminal end of a central parallel β-strand followed by a loop into the N-terminus of an α-helix, similar to BHPTP. Among these proteins, BHPTP shows especially strong structural homology with flavodoxin and Che Y. Both of these proteins have folding topologies of $1x,-2x,-1x,-1x$ for their central parallel β-sheets which closely resemble the topology $(1x,-2x,-1x)$ of BHPTP.

Comparison to PTP1B. Recently, the crystal structure of a truncated form (321 of 435 residues) of the protein tyrosine phosphatase PTP1B has been reported (Barford et al., 1994). PTP1B is a high molecular weight membrane-associated protein tyrosine phosphatase, exhibiting only limited sequence homology to BHPTP. PTP1B is much larger than BHPTP and contains a single catalytic phosphatase domain which consists of 8 α-helices, 12 β-strands, and several connecting loops. Despite the lack of sequence homology, there is local structural similarity between the two phosphatases (Figure 7). The hydrophobic core of both proteins consists of a four-stranded parallel β-sheet flanked by α-helices. Furthermore, the active site nucleophiles (C215 in PTP1B and C12 in BHPTP) are located at the C-terminal end of a central β-strand which is followed by a short loop into an α-helix. In PTP1B, the active site arginine, R221, is located three residues into the α-helix from its N-terminus and forms hydrogen bonds to a tungstate anion bound in the active site, coincident with the axis of the α-helix. In BHPTP, the location of the corresponding residue (R18) is ambiguous from the NMR data, although the presence of a proline residue at the Arg + 2 position suggests that $\alpha 1$ does not extend past R18 (Richardson & Richardson, 1988; Serrano et al., 1992; Blaber et al., 1993). Amino acid sequence alignment (Camici et al., 1989) reveals that other protein tyrosine phosphatases, i.e., cdc25-like PTPs, also have a proline residue at the Arg + 2 position. The location of the active site arginine near the N-terminus of an α-helix, as in BHPTP, as opposed to being well within the helix, as observed for PTP1B, places restrictions on the size of the binding site and may directly affect substrate binding preferences.

CONCLUSIONS

In this report, the first solution structure of a protein tyrosine phosphatase is presented. The structure is based upon over 1700 experimental restraints determined by multidimensional heteronuclear NMR methods. BHPTP consists of a central four-stranded parallel β-sheet flanked by four α-helices and one 3_{10} helix. The phosphate binding site is formed by residues at the C-terminus of $\beta 1$, the N-terminus of $\alpha 1$, and the loop connecting these two elements, based on ^1H and ^{15}N chemical shift changes observed in the presence and absence of phosphate and vanadate. Other residues involved in anion binding are found in $\alpha 1$, $\alpha 4$, $\alpha 5$, and some interconnecting loop regions.

To date, the structure of only one other PTP has been reported. That protein, PTP1B, belongs to a different class of PTPs and differs significantly from BHPTP. Despite the many differences between the two proteins, the central hydrophobic core and phosphate binding sites are structurally homologous. Comparisons of the structural motifs used in substrate recognition in these and other members of the PTP family will be important for elucidating the regulation and physiological roles of this important class of proteins. This structure will aid in the design of future mutagenesis experiments to probe the catalytic mechanism and substrate binding of this class of PTPs.

REFERENCES

Ab, E., Schuurman-Wolters, G. K., Saier, M. H., Reizer, J., Jacuinod, M., Roepstorff, P., Dijkstra, K., Scheek, R. M., & Robillard, G. T. (1994) *Protein Sci. 3*, 282–290.

Adams, M. J., Ford, G. C., Koekok, R., Lentz, P. J., Jr., McPherson, A., Jr., Rossmann, M. G., Smiley, I. E., Schevitz, R. W., & Wonnacott, A. (1970) *Nature 227*, 1098–1103.

Archer, S. J., Ikura, M., Torchia, D., & Bax, A. (1991) *J. Magn. Reson. 95*, 636–641.

Barford, D., Flint, A. J., & Tonks, N. K. (1994) *Science 263*, 1397–1404.

Bax, A., & Pochapsky, S. (1992) *J. Magn. Reson. 99*, 638–640.

Blaber, M., Zhang, X., & Matthews, B. W. (1993) *Science 260*, 1637–1640.

Bodenhausen, G., & Ruben, D. J. (1980) *Chem. Phys. Lett. 69*, 185–189.

Boivin, P., & Galand, C. (1986) *Biochem. Biophys. Res. Commun. 134*, 557–564.

Bränden, C.-I. (1980) *Q. Rev. Biophys. 13*, 317–338.

Brooks, B. R., et al. (1983) *J. Comput. Chem. 4*, 187–217.

Brünger, A. T. (1992) *X-PLOR Manual*, version 3.1, Yale University Press, Cambridge, MA.

Camici, G., Manao, G., Cappugi, G., Modesti, A., Stefani, M., & Ramponi, G. (1989) *J. Biol. Chem. 264*, 2560–2567.

Campbell, J. W., Watson, H. C., & Hodgson, G. I. (1974) *Nature 250*, 301–303.

Carson, M. (1987) *J. Mol. Graphics 5*, 103–106.

Charbonneau, H., & Tonks, N. K. (1992) *Annu. Rev. Cell. Biol. 8*, 463–493.

Chernoff, J., & Li, H. C. (1985) *Arch. Biochem. Biophys. 240*, 135–145.

Cho, H., Ramer, S. E., Itoh, M., Winker, D. G., Kitas, E., Bannwarth, W., Burn, P., Saito, H., & Walsh, C. (1991) *Biochemistry 30*, 6210–6216.

Cho, H., Krishnaraj, R., Kitas, E., Bannwarth, W., Walsh, C., & Anderson, K. S. (1992) *J. Am. Chem. Soc. 114*, 7296–7298.

Chui, D., Ong, C. J., Johnson, P., Teh, H.-S., & Marth, J. D. (1994) *EMBO J. 13*, 798–807.

Cirri, P., Chiarugi, P., Camici, G., Manao, G., Raugei, G., Cappugi, G., & Ramponi, G. (1993) *Eur. J. Biochem. 214*, 647–657.

Davis, J. P., Zhou, M.-M., & Van Etten, R. L. (1994a) *J. Biol. Chem. 269*, 8734–8740.

Davis, J. P., Zhou, M.-M., & Van Etten, R. L. (1994b) *Biochemistry 33*, 1278–1286.

Dissing, J., Johnsen, A. H., & Sensabaugh, G. F. (1991) *J. Biol. Chem. 266*, 20619–20625.

Dreusicke, D., Karplus, P. A., & Schulz, G. E. (1988) *J. Mol. Biol. 199*, 359–371.

Feng, G.-S., Hui, C.-C., & Pawson, T. (1993) *Science 259*, 1607–1611.

Fischer, E. H., Charbonneau, & Tonks, N. K. (1991) *Science 253*, 401–406.

Gilliland, G. L., & Quiocho, F. A. (1981) *J. Mol. Biol. 146*, 341–362.
Gu, M., York, J. D., Warshawsky, I., & Majerus, P. W. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 5867–5871.
Guan, K. L., & Dixon, J. E. (1991) *J. Biol. Chem. 266*, 17026–17030.
Hol, W. G. J., van Duijnen, P. T., & Berendsen, H. J. C. (1978) *Nature 273*, 443–446.
Hunter, T., & Cooper, J. A. (1985) *Annu. Rev. Biochem. 54*, 897–930.
Hyberts, S. G., Goldberg, M. S., Havel, T. F., & Wagner, G. (1992) *Protein Sci. 1*, 736–751.
Kay, L. E., & Bax, A. (1990) *J. Magn. Reson. 86*, 110–126.
Kay, L. E., Xu, G.-Y., Singer, A. U., Muhandiram, D. R., & Forman-Kay, J. D. (1993) *J. Magn. Reson. 101B*, 333–337.
Lazaruk, K. D. A., Dissing, J., & Sensabaugh, G. F. (1993) *Biochem. Biophys. Res. Commun. 196*, 440–446.
Manao, G., Pazzagi, L., Cirri, P., Caselli, A. G., Cappugi, G., Saeed, A., & Ramponi, G. (1992) *J. Protein Chem. 11*, 333–345.
Marion, D., Ikura, M., Tschudin, R., & Bax, A. (1989a) *J. Magn. Reson. 85*, 393–399.
Marion, D., Ikura, M., & Bax, A. (1989b) *J. Magn. Reson. 84*, 425–430.
Messerle, B. A., Wider, G., Otting, G., Weber, C., & Würthrich, K. (1989) *J. Magn. Reson. 85*, 608–613.
Neri, D., Szyperski, T., Otting, G., Senn, H., & Wüthrich, K. (1989) *Biochemistry 28*, 7510–7516.
Nilges, M., Clore, G. M., & Gronenborn, A. M. (1988) *FEBS Lett. 239*, 129–136.
Olejniczak, E. T., & Eaton, H. L. (1990) *J. Magn. Reson. 87*, 628–632.
Ploegman, J. H., Drent, G., Kalk, K. H., & Hol, W. G. J. (1978) *J. Mol. Biol. 123*, 557–594.
Ramponi, G., Manao, G., Camici, G., Cuppugi, G., Ruggiero, M., & Bottaro, D. P. (1989) *FEBS Lett. 250*, 467–473.
Rao, S., & Rossmann, M. (1973) *J. Mol. Biol. 76*, 241–256.
Richardson, J. S., & Richardson, D. C. (1988) *Science 240*, 1648–1652.
Rohan, P. J., Davis, P., Moskaluk, C. A., Kearns, M., Krutzsch, H., Siebenlist, U., & Kelly, K. (1993) *Science 249*, 1763–1766.
Rucker, S. P., & Shaka, A. J. (1989) *Mol. Phys. 68*, 509–517.
Ruggiero, M., Pazzagli, C., Rigacci, S., Magnelli, Raugei, G., Berti, A., Chiarugi, V. P., Pierce, J. H., Camici, G., & Ramponi, G. (1993) *FEBS Lett. 326*, 294–298.
Saini, M. S., Buchwald, S. C., Van Etten, R. L., & Knowles, J. R. (1981) *J. Biol. Chem. 256*, 10453–10455.
Serrano, L., Sancho, J., Hirshberg, M., & Fersht, A. R. (1992) *J. Mol. Biol. 227*, 544–559.
Shen, S.-H., Bastien, L., Posner, B. I., & Chretien, P. (1991) *Nature 352*, 763–739.
Smith, W. W., Burnett, R. M., Darling, G. D., & Ludwig, M. L. (1977) *J. Mol. Biol. 117*, 195–225.
Sternberg, M., & Thornton, J. (1977) *J. Mol. Biol. 110*, 269–283.
Stock, A. M., Martinez-Hackert, E., Rasmussen, B. F., West, A. M., Stock, J. B., Ringe, D., & Petsko, G. A. (1993) *Biochemistry 32*, 13375–13380.
Stockman, B. J., Krezel, A. M., Markley, J. M., Leonhardt, K. G., & Straus, N. A. (1990) *Biochemistry 29*, 9600–9609.
Taga, E. M., & Van Etten, R. L. (1982) *Arch. Biochem. Biophys. 214*, 505–515.
Tonks, N. K., Diltz, C. D., & Fischer, E. H. (1988) *J. Biol. Chem. 263*, 6731–6737.
Van Etten, R. L., Waymack, P. P., & Rehkop, D. M. (1974) *J. Am. Chem. Soc. 96*, 6782–6785.
Vuister, G. W., & Bax, A. (1993) *J. Am. Chem. Soc. 115*, 7772–7777.
Vuister, G. W., Clore, G. M., Gronenborn, A. M., Powers, R., Garrett, D. S., Tschudin, R., & Bax, A. (1993) *J. Magn. Reson. 101B*, 210–213.
Waheed, A., Laider, P. M., Wo, Y.-Y. P., & Van Etten, R. L. (1988) *Biochemistry 27*, 4265–4273.
Walton, K. M., & Dixon, J. E. (1993) *Annu. Rev. Biochem. 62*, 101–120.
Wo, Y.-Y. P., Zhou, M.-M., Stevis, P., Davis, J. P., Zhong, Z.-Y., & Van Etten, R. L. (1992a) *Biochemistry 31*, 1712–1721.
Wo, Y.-Y. P., McCormack, A. L., Shabanwitz, J., Hunt, D. F., Davis, J. P., Mitchell, G. L., & Van Etten, R. L. (1992b) *J. Biol. Chem. 267*, 10856–10865.
Wüthrich, K., Billeter, M., & Braun, W. (1983) *J. Mol. Biol. 169*, 949–961.
Yang, Q., & Tonks, N. K. (1991) *Cell 67*, 661–673.
Yarden, Y., & Ullrich, A. (1988) *Annu. Rev. Biochem. 57*, 443–478.
Yi, T., Cleveland, J. L., & Ihle, J. N. (1992) *Mol. Cell. Biol. 12*, 836–846.
Zhang, Z.-Y., & Van Etten, R. L. (1990) *Arch. Biochem. Biophys. 228*, 39–49.
Zhang, Z.-Y., & Van Etten, R. L. (1991) *J. Biol. Chem. 266*, 1516–1525.
Zhang, Z.-Y., & Dixon, J. E. (1994) *Adv. Enzymol. 68*, 1–36.
Zhang, Z.-Y., Davis, J. P., & Van Etten, R. L. (1992) *Biochemistry 31*, 1701–1711.
Zhou, M.-M. (1993) Ph.D. Thesis, Purdue University, West Lafayette, IN.
Zhou, M.-M., Davis, J. P., & Van Etten, R. L. (1993) *Biochemistry 32*, 8479–8486.
Zhou, M.-M., Logan, T. M., Theriault, Y., Van Etten, R. L., & Fesik, S. W. (1994) *Biochemistry 33*, 5221–5229.